Available online at www.sciencedirect.com

ScienceDirect

Procedia - Social and Behavioral Sciences 160 (2014) 92 – 101

Procedia
Social and Behavioral Sciences

XI Congreso de Ingenieria del Transporte (CIT 2014)

Importance of Dwelling and Neighbourhood Attributes in Residential Location Modelling: Best Worst Scaling vs. Discrete Choice

Camila Balbontin[a,*] , Juan de Dios Ortúzar[a], Joffre D. Swait[b]

[a]Pontificia Universidad Católica de Chile, Vicuña Mackenna 4860, Santiago, Chile

[b]University of South Australia, Level 13, 140 Arthur Street, North Sydney, Australia

Abstract

Traditional discrete choice experiments do not differentiate between the intrinsic importance of an attribute and that associated with its levels of variation. It has been suggested recently that best-worst (B-W) scaling (Case 2) allows for this differentiation. Here we pool B-W answers with binary stated choice (SC) data to study the importance of dwelling and neighbourhood attributes for apartment seekers in the centre of Santiago, Chile. Previous research had shown how these diverse elicitation methods can be pooled (albeit without including any type of heterogeneity), suggesting that the "best" (as opposed to "worst") responses are most compatible with the binary SC data. In this paper we extend this work to allow for heterogeneity in preferences through (a) systematic taste variations alone, (b) correct treatment of panel effects alone, and finally, (c) the combined effect of both. In all cases the best resulting model is obtained by pooling the "best" answers with the binary SC, under the assumption of common and specific attributes to each dataset. Nevertheless, when the model included only unobserved heterogeneity through error components (to treat the *panel effect*), the datasets did not pool as well as when we did not consider it. The joint model had half of the attributes specific to each dataset while in the previous case only two were specific. On the other hand, when considering observed and unobserved heterogeneity, the two datasets pooled better than in the other cases, needing to consider only one attribute as specific to each set (and the remaining seven as common). We also analysed how the inclusion of heterogeneity changed the attribute estimates. In particular, the scales that can be constructed for the attributes and for their levels of variation do not change their relative order, but the magnitudes do, especially in cases where the estimates were not significantly different from zero at the 95% confidence level.

* Correspondence author: Tel: +56-2-2354 4270: E-mail: cami.balbontin.t@gmail.com

Peer-review under responsibility of CIT 2014.

Keywords: discrete choice experiments; residential location; attribute importance; best-worst scaling; heterogeneity

Peer-review under responsibility of CIT 2014.

doi:10.1016/j.sbspro.2014.12.120

1. Introduction

Although traditional stated choice (SC) experiments have been applied to study individual choices in many areas they have several limitations. One of the most discussed is their inability to differentiate between the impacts of the attributes by themselves and that associated with their levels of variation. Best-Worst Scaling (Case 2) allows for this differentiation (Louviere and Swait, 1997).

We combine binary stated choice (SC) answers with best-worst (B-W) responses in order to study the importance of several dwelling and neighbourhood attributes for apartment seekers in the centre of Santiago, Chile. These different elicitation methods were pooled by Balbontin et al. (2013) without allowing for interpersonal preference heterogeneity. Thus, we propose to extend their work allowing for both observed heterogeneity in preferences (i.e. systematic taste variations) and unobserved heterogeneity. The former will be included by letting the experimental attributes interact with the respondents' socioeconomic characteristics (Ortúzar and Willumsen, 2011, p. 279) whereas the latter will be included through error components in a simple mixed logit (ML) structure (Train, 2009; Hess and Rose, 2012) to allow for the correct treatment of the repeated observations' nature of the data.

The data used was collected as part of a study by Torres et al. (2013) who surveyed people who intended to rent an apartment in the centre of Santiago with the objective of studying the importance of several dwelling and neighbourhood attributes in location choice. In this case, each respondent was asked to provide three different responses to the survey: (1) which of the attributes (level) presented in each profile was the most attractive ('best' answer); (2) which one was the least attractive ('worst' answer), and (3) if the respondent would or would not buy/rent the apartment shown in the profile (stated choice, SC, response).

The objective of this paper is to study how the data sets can be pooled while allowing for each source of preference heterogeneity individually and jointly. Likewise, we will see how this inclusion changes the estimates and the relative scale that can be constructed for the attributes and for their levels of variation.

2. Design of the Survey

The survey was answered by 203 people who planned to rent or buy an apartment in the centre of Santiago. Each of them responded to eight different apartment descriptions (see Figure 1). The apartment profiles were defined by an experimental design based on eight different attributes that described both the neighbourhood in which the apartment was located and the dwelling itself. As mentioned before, in each profile the person provided three different responses: (a) the attribute level that seemed most attractive ('best' answers), (b) the attribute level that seemed least attractive ('worst' answers), and (c) if s/he would or would not consider acquiring the apartment profile shown (binary SC answers). An example of the basic survey task can be found in Figure 1.

The 203 individuals surveyed answered in full the binary SC question (thus, there are 1,624 observations of the SC response), but not all of them answered fully the 'best' and 'worst' questions in all eight tasks. Hence, there are 1,577 'best' answers and 1,514 'worst' answers.

The data was coded using 'effects code' (Louviere and Swait, 1997), that is, the highest level of each attribute was coded as +1 and the lower as -1 (see Table 1). Seven of the eight attributes had two levels, and the attribute *Rent* had eight different levels. However, each individual was presented with only two of the levels depending on the number of rooms the person required in the apartment. With the objective of having equal magnitudes for the slopes and intercepts of all attributes, the attribute *Rent* was coded normalizing the levels between +1 and -1.

3. Best-Worst versus Stated Choice Utility Functions

We used logit models to estimate the parameters in the utility functions for the 'best', 'worst' and/or binary SC responses. As mentioned in section 2, the 'best' and 'worst' answers are among attribute levels.

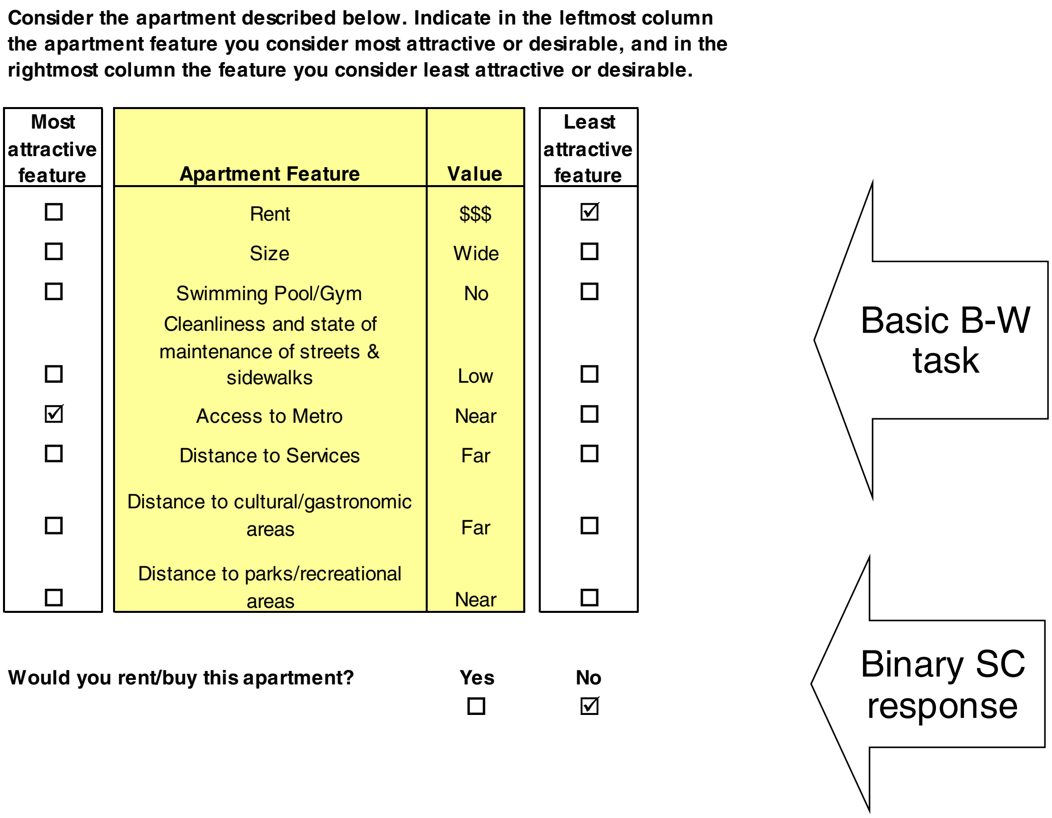

Figure 1: Example of Profile shown in the Survey

Thus, for each attribute we were able to estimate a utility function containing an *intercept* and a *slope*. The intercept represents the impact that the attribute has in the utility function by itself, and the slope the impact of the difference in the levels shown (Louviere and Swait, 1997). This is due to the way in which the attribute levels were coded (as discussed above). Also, it should be noted that the utility functions for the 'best' and 'worst' data are not the same. In general, it is assumed that the utility function for the 'worst' answers is the inverse of that corresponding to the 'best' answers (Louviere and Swait, 1997). We specified the utility functions as inverse but when we estimated the models jointly included a scale factor as usual (Louviere *et al.*, 2000). On the other hand, the binary SC responses are based on the whole profile. For this reason, the utility function associated with the answer 'I would rent the apartment' includes all the attribute levels shown in the profile (i.e. the utility for its binary alternative is specified as zero), and for each one we are able to estimate its slope.

When pooling the answers from each data set ('best', 'worst' and/or SC) we can determine if certain attribute level slopes can be considered common (i.e. with the same parameter). Again, when mixing data of a different nature we include appropriate scale factors (as, in principle, the standard deviation of the stochastic errors in the utility functions may vary in each data set, Ortúzar and Willumsen, 2011, Chapter 7). Thus, analytically, the utility functions estimated for each data set were specified as follows:

Best Data:

$U_{Metro} = \lambda_{Best} \cdot (C_{Metro\_Best} + \theta_{Metro\_Best} \cdot X_{Metro}) + \varepsilon_{Best}$

$U_{Parks} = \lambda_{Best} \cdot (C_{Parks\_Best} + \theta_{Parks\_Best} \cdot X_{Parks}) + \varepsilon_{Best}$

and so on for each attribute.

Worst Data:

$U_{Metro} = \lambda_{Worst} \cdot (- C_{Metro\_Worst} + \theta_{Metro\_Worst} \cdot (-X_{Metro})) + \varepsilon_{Worst}$

$U_{Parks} = \lambda_{Worst} \cdot (- C_{Parks\_Worst} + \theta_{Parks\_Worst} \cdot (-X_{Parks})) + \varepsilon_{Worst}$

and so on for each attribute.

SP Data:

$U_{I\ would\ rent\ the\ apartment} = \lambda_{SC} \cdot (C_{Yes} + \theta_{Metro\_SC} \cdot X_{metro} + \theta_{Parks\_SC} \cdot X_{Parks} + \dots) + \varepsilon_{SC}$

$U_{I\ would\ not\ rent\ the\ apartment} = 0 + \varepsilon_{SC}$

The inclusion of the scale factor allows the errors to distribute as follows:

ε_{Best} ~ IID Gumbel $(0, \lambda_{Best})$

ε_{Worst} ~ IID Gumbel $(0, \lambda_{Worst})$

ε_{SC} ~ IID Gumbel $(0, \lambda_{SC})$

When estimating a model pooling the datasets we examine which slopes of the attributes can be considered common. Thereby, testable restrictions are added as follows:

$\theta_{Metro\_Best} = \theta_{Metro\_Worst} = \theta_{Metro\_SC}$

$\theta_{Parks\_Best} = \theta_{Parks\_Worst} = \theta_{Parks\_SC}$

and so on for each attribute.

4. Inclusion of Heterogeneity

4.1. Observed Heterogeneity (STV)

We considered including observed heterogeneity though systematic taste variations (STV). Torres et al. (2013) asked each individual for certain socioeconomic characteristics. Thus, we were able to analyse which of these had significant and meaningful interactions with the slopes of the attributes. In this way we could estimate individual preferences depending on their socioeconomic characteristics.

4.2. Unobserved Heterogeneity (PD)

We included unobserved heterogeneity through error components to treat correctly the pseudo panel data (PD) nature of our data. As individuals were exposed to eight different scenarios, it cannot be assumed that their responses are independent; in classic SC models, including this effect has been shown to have a significant impact in the model's goodness-of-fit (Ortúzar and Willumsen, 2011, Chapter 8). In our models, this *panel data effect* was considered by adding an error component to the utility function that distributes $N(0, \sigma^2_q)$; its presence in the utility function permits within individual correlation across alternatives.

5. Pooling the Data Sets

The main objective of this paper was to study how the data sets could be pooled while adding heterogeneity. Thus, we studied how the models without and with heterogeneity behaved using the different combinations of the three types of answers ('best', 'worst' and SC) available.

Table 1: Definition and Coding of Attributes

Attribute	Levels	Code
Access to Metro	Far, more than five blocks away	+1
	Near, less than five blocks away	-1
Distance to parks/recreational areas	Far, more than five blocks away	+1
	Near, less than five blocks away	-1
Rent	Ch$ 120,000	-1
	Ch$ 140,000	-0.819
	Ch$ 171,000	-0.538
	Ch$ 199,000	-0.285
	Ch$ 217,000	-0.122
	Ch$ 253,000	+0.204
	Ch$ 293,000	+0.567
	Ch$ 341,000	+1
Swimming Pool/Gym	Yes	+1
	No	-1
Cleanliness and state of maintenance of streets & sidewalks	Good	+1
	Poor	-1
Distance to Services	Far, more than five blocks away	+1
	Near, less than five blocks away	-1
Distance to cultural/gastronomic areas	Far, more than five blocks away	+1
	Near, less than five blocks away	-1
Size	Wide	+1
	Normal	-1

(*) At the time of the survey 1 US$ = 500 Ch$

5.1. Best+SC Models

First, we used only the 'best' and SC answers to estimate the different models. The log-likelihood estimates for the different models, as well as the number of estimated parameters and the number of attributes that could be considered common between the two types of answers are shown in Table 2.

To check if the inclusion of heterogeneity was necessary we applied the Likelihood Ratio (LR) test (Ortúzar and Willumsen, 2011, p. 281), as follows:

$$LR = -2 * (-3{,}161 + 2{,}835) = 652 > \chi^2_{31;0.01} = 61.1$$

where the general model is the one including both types of heterogeneity (with a log-likelihood at convergence of -2,835), and the restricted model is the original one without heterogeneity (log-likelihood = -3,161). As LR is substantially larger than the critical $\chi 2$ value for 31 degrees of freedom (i.e. the additional number of parameters of the general model) we rejected the null hypothesis that the models are equivalent at the 99% level. This result undergirds our conclusion that adding both types of heterogeneity indeed improves the pooled model.

Table 2: *Best+SC* Models

	Log-likelihood	Number of estimated parameters	Number of common attributes among answers
Without Heterogeneity	-3,161	19	6
With PD	-2,917	23	4
With STV	-3,057	50	5
With STV and PD jointly	-2,835	50	7

It is interesting also to notice that when including STV and PD heterogeneity the 'best' and SC answers appeared to reflect similar underlying evaluations, considering only one of the attributes as specific to each data set. The attribute that was considered specific was the *Rent*, which seemed to have a different behaviour in each data set.

5.2. Worst+SC Models

When considering the 'worst' with the SC answers, we found that when including heterogeneity by treating the *panel data effect* (by itself and including it jointly with STV heterogeneity), the 'worst' answers could not be pooled with the SC answers. This means that when analysing which attributes could be considered common among the answers we found that, in fact, all attributes should be considered as specific to the individual elicitation method.

It is important to mention that the inclusion of both types of heterogeneity improved both the 'worst' and SC model by themselves. Consequently, this means that the 'worst' answers do not behave in a similar way to the SC elicitation. The fact that they could be pooled while not including PD heterogeneity suggests that the modelling of heterogeneity is necessary in data pooling efforts; its omission may lead to major inference errors.

5.3. Best+Worst+SC Models

Finally, we attempted to estimate models including the three answers ('best', 'worst', and SC) including and not including heterogeneity. The 'worst' answers could not consider any attributes common with neither of the other answers. We were only able to estimate models without heterogeneity, and when including only STV heterogeneity. The results are shown in Table 3.

Table 3: *Best+Worst+SC* Models

	Log-likelihood	Number of estimated parameters	Number of common attributes among answers
Without Heterogeneity	-5,323	29	4
With STV	-5,188	64	5

We can analyse if the inclusion of STV heterogeneity was necessary applying the Likelihood Ratio (LR) test again:

$$LR = -2 * (-5{,}323 + 5{,}188) = 270 > \chi^2_{35;0.01} = 66.6$$

and this shows that the inclusion of STV heterogeneity is necessary and brings relevant information to the model. Likewise, we can also see in the fourth column of Table 3 that when including STV heterogeneity the answers adjusted better than when not including it, allowing to consider more attributes as common among the different datasets.

6. Analysis of the Inclusion of Heterogeneity

The 'best' and SC answers yield the best resulting model. These elicitation methods were the only ones that could be pooled while including both types of heterogeneity, which improved significantly the model. Thus, we are now able to analyse how this inclusion changed the estimates for each attribute.

6.1. Intercepts

As explained previously, the intercepts for each attribute (see Table 4) represent their intrinsic impact in the utility function. In this case, these impacts are relative to the impact of the attribute *Distance to services*, which was considered as base (i.e. was not estimated and considered to be equal to zero).

Table 4: Estimated Intercepts

Attribute / Intercept	**Without Heterogeneity**	**PD+STV**
Access to Metro	0.269 (4.5)	0.226 (4.0)
Distance to parks/recreational areas	-0.085 (1.3)	-0.055 (0.9)
Rent	0.044 (0.8)	-0.059 (0.9)
Swimming Pool/Gym	-0.177 (2.8)	-0.265 (3.6)
Cleanliness and state of maintenance of streets & sidewalks	-0.273 (4.0)	-0.386 (4.8)
Distance to services	0	0
Distance to cultural gastronomic areas	-0.086 (1.5)	-0.177 (2.7)
Size	0.504 (7.3)	0.508 (8.0)

As can be seen in Table 4, the inclusion of heterogeneity does not change significantly the relative position of the attributes nor their value. Nevertheless, the two intercepts that are not significantly different from zero with an 80% of confidence level in the PD+STV model do change their relative position. These correspond to the attributes *Distance to parks/recreational areas* and *Rent*. Moreover, the intercept for the attribute *Rent* also changes its sign. This could be expected, because these two intercepts do not appear to have a significant impact on the utility function (relative to the attribute *Distance to services*), thus, their values do not provide meaningful information. In terms of the value of the slopes, these are also the two that appear to have the largest difference. The rest of the attributes seem to have the same relative position.

6.2. Slopes

The slopes of the attributes represent the impact that the variation in the levels shown in the survey have in the utility function.
Table 5 show the slopes estimated for each attribute (when the attribute was considered specific to each dataset, both slopes are shown). The slopes estimated for the PD+STV model correspond to the average calculated over each individual in the dataset. In this case, we can see that the signs of the slopes for each attribute are equal for both models. Also, all slopes for the model without heterogeneity are significantly different from zero at the 95% confidence level, except for that of the attribute *Distance to parks/recreational areas* estimated for the SC answers.

The relative positions of the slopes do not vary as much as the values of them do. For instance, the slope of the attribute *Rent* in both models has the largest impact, however, its value varies 171% for the slope estimated for the 'best' answers in the PD+STV model and 203% for the SC answers. This is the slope that increases the most when including heterogeneity. Thus, with this inclusion the attribute *Rent* has a larger impact in the utility function in comparison to the other attributes, having a considerably larger slope than the attribute that follows.

Table 5: Estimated Slopes

Attribute / Slope	Without Heterogeneity	PD+STV
Access to Metro	-0.413 (10.0)	-0.514
Distance to parks/recreational areas	Best: -0.359 (5.7) SC: -0.087 (1.6)	-0.296
Rent	-0.472 (6.3)	Best: -1.279 SC: -1.429
Swimming Pool/Gym	0.115 (3.2)	0.121
Cleanliness and state of maintenance of streets & sidewalks	0.324 (7.3)	0.394
Distance to Services	Best: -0.375 (6.5) SC: -0.186 (3.5)	-0.364
Distance to cultural/gastronomic areas	-0.149 (4.5)	-0.181
Size	0.299 (9.4)	0.384

The two attributes that were considered specific to each data set in the model that did not include heterogeneity were *Distance to services* and *Distance to parks/recreational areas.* In both cases, the slope that appeared to be more similar to the one estimated in the PD+STV model was that estimated using the 'best' answers. Leaving aside these attributes, the third attribute that had the largest impact was *Cleanliness and state of maintenance of streets & sidewalks*, followed by *Size* and finally by *Swimming pool/gym.*

In conclusion, the relative position of the attribute slopes did not change significantly when including heterogeneity. Also, this inclusion increased the impact of all attribute levels, except when comparing them with the slopes estimated without heterogeneity for *Distance to services* and *Distance to parks/recreational areas* using the 'best' answers (which were slightly larger).

7. Analysis of the Model

The inclusion of heterogeneity appeared to be necessary, with the best-pooled (PD+STV) model. It is interesting to note how using B-W data provides more information about the attributes, allowing to differentiate between the impact that the attribute has by itself (relative to *Distance to services*) and the impact that the attribute levels have in the utility function. The attributes with high impact in the utility function both by themselves and because of the variation in their levels are *Cleanliness and state of maintenance of streets & sidewalks* and *Size*. These two are followed by *Access to Metro*, that has a medium relative impact by itself (relative to *Distance to services*), but a high impact due to the variation in levels. This means that if the variation in levels had been smaller, this attribute would have had a smaller impact on the utility function. The attribute *Distance to services* had a high impact due to the variation in levels, and its impact by itself was not estimated (it was the base). The attributes *Distance to cultural/gastronomic areas* and *Swimming pool/gym* both had a medium impact by themselves (relative to *Distance to services*) and a small impact due to the variation in levels shown. The attribute *Distance to parks/recreational areas* seemed to have a small impact by itself and a medium impact due to the variation of levels shown.

The most interesting attribute was the *Rent*, which appeared to have the largest impact due to the variation of levels shown in the survey, but had a small impact by itself (relative to *Distance to services*). This differentiation that can be obtained through the B-W task may be relevant in decision making. For example, if a real-estate developer would like to rent more apartments but was not able to make this differentiation, he could decide to decrease the rent expecting a big increase in the number of potential tenants. But if this decrease translates into a smaller variation in the rents of the apartments in the sector, this attribute would not be as significant to the potential tenants as the original model showed. Thus, the increase in the number of tenants would not be as high as the real-estate developer was expecting. Hence, using B-W increases understanding on the behaviour of the attributes, providing information that could affect substantially managerial decisions concerning marketing actions.

8. Conclusions

The objective of this paper was to analyse how the inclusion of heterogeneity affected the pooling of three different data sets, one corresponding to 'best' answers, one to 'worst' answers and one to SC answers. We found that the 'best' and SC answers could be pooled when including heterogeneity through (a) systematic taste variations (STV), (b) through error components to treat correctly the panel effect (PD) and (c) when combining both types of effects. The latter improved the model significantly, demonstrating the necessity of including both types of heterogeneity. It is also important to mention that when not including heterogeneity, the 'best' and SC answers considered four attributes as specific to each environment, and when including PD+STV only two attributes were required to be specific. Hence, the inclusion of PD and STV allowed for a better fit of the 'best' and SC answers. On the other hand, the 'worst' answers did not behave similar to the 'best' nor to the SC answers when including PD and could not consider any attribute common among them. This shows the importance of including heterogeneity, as the 'worst' answers could be pooled with the other two when not including heterogeneity.

The relative positions of the intercepts and slopes estimated for each model did not change significantly when including heterogeneity, although their values changed, especially in the case of the slopes. The slope for the attribute *Rent* gained the largest impact when including heterogeneity. Nevertheless, all slopes increased their value when including heterogeneity, except for those attributes that were considered specific to each dataset when not including heterogeneity. In this case, two attributes were considered specific (*Distance to services* and *Distance to parks/recreational areas*) and the slopes estimated using the 'best' answers were the ones more similar to those estimated when including heterogeneity, being larger the ones estimated in the first model.

The preferred model was estimated using the 'best' and SC answers and including both types of heterogeneity (PD+STV). This model showed that the most relevant attributes, both by themselves (relative to *Distance to services*) and due to the variation in levels shown in the survey were *Cleanliness and state of maintenance of streets & sidewalks, Size*, and *Access to Metro.* Those that appeared less relevant were *Swimming pool/Gym* and *Distance to cultural/gastronomic areas*. The attribute *Distance to services* had a large impact due to the variation in the levels shown. Finally, the *Rent* showed a high impact in the utility function due to the variation in levels shown, but a small impact by itself.

In conclusion, the detailed analysis made possible for each attribute suggests that the joint use of B-W Scaling (Case 2) with traditional discrete choice experiments allows for a better understanding of housing choices.

Acknowledgements

We would like to thank the support of the National Fund for Scientific and Technological Research (Fondecyt Project 1100068), the Millennium Institute on Complex Engineering Systems (ICM: P05-004F; Fondecyt: FB016), and the Centre for Urban Sustainable Development CEDEUS (Conicyt/Fondap/15110020) to this investigation. We are also grateful to Ignacia Torres for having collected the data.

References

Balbontin, C., Ortúzar, J. de D. and Swait, J.D. (2013) Importancia de los atributos de vivienda y barrio en localización residencial: una aplicación del método *Best-Worst* al centro de Santiago. *XVI Congreso Chileno de Ingeniería de Transporte,* Universidad de Chile, Santiago (in Spanish).

Coast, J., Flynn, T.N., Natarajan, L., Sproston, K., Lewis, J., Louviere, J.J. and Peters, T.J. (2008) Valuing the ICECAP capability index for older people. *Social Science and Medicine* **67**, 874-82.

Couzner, L., Ratcliffe, J., Lester, L., Flynn, T. and Crotty, M. (2013) Measuring and valuing quality of life for public health research: application of the ICECAP-O capability index in the Australian general population. *International Journal of Public Health* **58**, 367-376.

Finn, A. and Louviere, J.J. (1992) Determining the appropriate response to evidence of public concern: the case of food safety. *Journal of Public Policy and Marketing* **11**, 12–25.

Flynn, T.N. (2010) Using conjoint analysis and choice experiments to estimate quality adjusted life year values: issues to consider. *Pharmacoeconomics* **28**, 711-722.

Flynn, T.N., Louviere, J.J., Peters, T.J. and Coast, J. (2010) Using discrete choice experiments to investigate heterogeneity in preferences for quality of life: variance scale heterogeneity matters. *Social Science and Medicine* **70**, 1957-1965.

Flynn, T.N., Peters, T.J. and Coast, J. (2013) Quantifying response shift or adaptation effects in quality of life by synthesizing best-worst scaling and discrete choice data. *Journal of Choice Modelling* **6**, 34-43.

Greene, M. and Ortúzar, J. de D. (2002). Willingness to pay for social housing attributes: a case study from Chile. *International Planning Studies* **7**, 55–87.

Hess, S. and Rose, J.M. (2012) Can scale and coefficient heterogeneity be separated in random coefficients models? *Transportation* **39**, 1225-1239.

Louviere, J.J. (2013) Modelling single individuals: the journey from psych lab to the app store. In S. Hess and A.J. Daly (eds.), *Choice Modelling: The State of the Art and the State of Practice*. Edward Elgar, Cheltenham.

Louviere, J.J. and Swait, J.D. (1997) Separating weights and scale values in conjoint tasks using choices of best and worst attribute levels. Working Paper, Centre for the Study of Choice, University of Technology Sydney.

Louviere, J.J., Hensher, D.A. and Swait, J.D. (2000) *Stated Choice Methods: Analysis and Applications*. Cambridge University Press, Cambridge.

Louviere, J.J., Street, D., Burgess, L., Wasi, N., Islam, T. and Marley, A.A. (2008) Modelling the choices of individual decision-makers by combining efficient choice experiment designs with extra preference information. *Journal of Choice Modelling* **1**, 128-163.

Marley, A.A. and Louviere, J.J. (2005) Some probabilistic models of best, worst, and best–worst choices. *Journal of Mathematical Psychology* **49**, 464-480.

McCullagh, P. and Nelder, J.A. (1989) *Generalized Linear Models*. Second Edition. Chapman and Hall, London.

McFadden, D. (1974) The measurement of urban travel demand. *Journal of Public Economics* **3**, 303–328.

Munizaga, M. and Alvarez-Daziano, R. (2002) Evaluation of mixed logit as a practical modelling alternative. *Proceedings European Transport Conference*, Cambridge.

Ortúzar, J. de D., Martínez, F.J. and Varela, F.J. (2000) Stated preferences in modelling accessibility. *International Planning Studies* **5**, 65–85.

Ortúzar, J. de D. and Willumsen, L.G. (2011) *Modelling Transport*. Fourth Edition, John Wiley and Sons, Chichester.

Swait, J.D. and Bernardino, A. (2000) Distinguishing taste variation from error structure in discrete choice data. *Transportation Research* **34B**, 1-15.

Swait, J.D. and Louviere, J.J. (1993) The role of the scale parameter in the estimation and comparison of multinomial logit models. *Journal of Marketing Research* **30**, 305-314.

Szeinbach, S.L., Barnes, J.H., McGhan, W.F., Murawski, M.M. and Corey, R. (1999) Using conjoint analysis to evaluate health state preferences. *Drug Information Journal* **33**, 849-858.

Torres, I., Greene, M. and Ortúzar, J. de D. (2013) Valuation of housing and neighbourhood attributes for city centre location: a case study in Santiago. *Habitat International* **39**, 62-74.

Train, K.E. (2009) *Discrete Choice Methods with Simulation.* Second Edition, Cambridge University Press, Cambridge.